Contextual Time Series Change Detection

Xi C. Chen* Karsten Steinhaeuser* Shyam Boriah* Snigdhansu Chatterjee† Vipin Kumar*

Abstract

Time series data are common in a variety of fields ranging from economics to medicine and manufacturing. As a result, time series analysis and modeling has become an active research area in statistics and data mining. In this paper, we focus on a type of change we call *contextual time series change* (CTC) and propose a novel two-stage algorithm to address it. In contrast to traditional change detection methods, which consider each time series separately, CTC is defined as a change relative to the behavior of a group of related time series. As a result, our proposed method is able to identify novel types of changes not found by other algorithms. We demonstrate the unique capabilities of our approach with several case studies on real-world datasets from the financial and Earth science domains.

1 Introduction

Time series data is ubiquitous in a wide range of applications from financial markets to manufacturing, from health care to the Earth sciences, and many others. As a result, time series analysis and modeling has become an active area of research in statistics and data mining [1, 11, 14]. Of particular interest within this realm are the problems of change detection [3, 6, 8, 9]. In this paper, we focus on a type of change we call *contextual time series change* (CTC) and present a novel approach for addressing the CTC detection problem. Below we give some background on change detection along with an intuitive description of a *contextual* change using a simple illustrative example; a formal definition of the problem follows in Section 2.

Traditional time series change detection is defined with respect to values in some portion of the same time series. By contrast, *contextual change* refers to a deviation in behavior of an object (the *target* time series) with respect to its *context*. The context consists of a collection of time series that exhibit similar behavior to the target time series for some period of time.

Simply put then, a contextual change can be described as a target time series behaving similarly to the related series for some period of time but then diverging from them. Fig. 1 and Fig. 2 illustrate several examples of contextually changed and unchanged time series. In both of the figures, the black line indicates the target time series and gray lines show the related time series. Fig. 1 shows three types of contextually changed time series. In Type 1, the target series changes abruptly around time step $t = 100$ while the context remains relatively stable; in Type 2, the context exhibits a collective change whereas the target series remains stable; and in Type 3, although both the target time series and its context keep changing during the whole period, after $t = 100$, their behaviors are no longer similar – that is, they begin to diverge. The last two types of changes are uniquely addressed by CTC detection and cannot be found by traditional time series change detection algorithms. For comparison, Fig. 2 shows two types of contextually unchanged time series. In the top panel, both the target time series and its context are stable while in the bottom panel, both of them change similarly during the whole period. Since the target time series does not diverge from its context in ether of these scenarios, they are considered to be contextually unchanged.

CTC detection is useful in many real-world settings. For example, consider a collection of sensors that measure the temperature in a factory. It is very likely that the sensors exhibit a diurnal pattern, with temperatures increasing during the morning hours and cooling in the evenings. Thus, a change in the sensor readings does not necessarily constitute an event. However, if a single sensor does *not* change similarly to the others, it may indicate a sensor failure or an unusual condition in the sensor environment. Similar situations arise in a wide range of application settings.

The **key contributions** of this paper can be summarized as follows:

1. We provide a formal definition of the problem we call ***contextual time series change*** (CTC) detection, which is distinct from traditional time series change detection and is not properly addressed by existing approaches.

2. We propose a new similarity function called k^{th}

*Department of Computer Science & Engineering, University of Minnesota. {chen,ksteinha,sboriah,kumar}@cs.umn.edu

†School of Statistics, University of Minnesota. chatterjee@stat.umn.edu

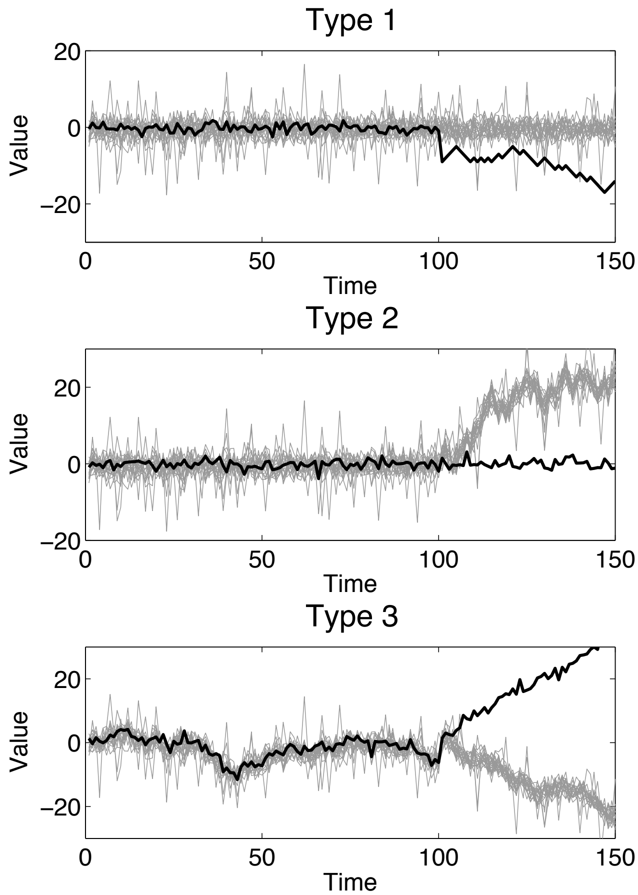


Figure 1: Illustration of different types of contextual changes in time series. In general, the target time series (black) behaves differently from the context after $t = 100$ in all the scenarios. Type 2 and Type 3 are uniquely addressed by CTC change detection algorithms.

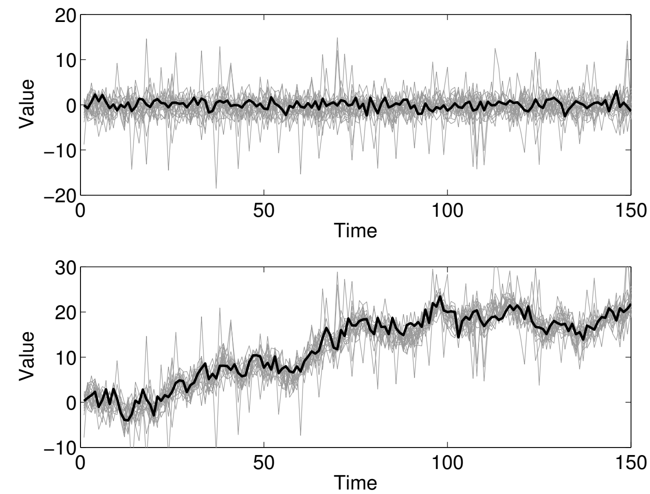


Figure 2: Illustration of contextually unchanged time series. In both of the scenarios, the target time series (black) behaves similarly to their context (grey) during the entire period. Note that the target time series in the lower panel may be considered as changed by traditional change detection schemes.

order statistic distance. We also provide a method to estimate the best k assuming that the probability of a time step is an outlier can be estimated from domain knowledge.

3. We derive a new metric called ***time series area depth*** to measure the deviation of the target series from the context group members. Time series area depth–which is anti-correlated to the probability that the target object belongs to a group–is a good indicator of contextual change.

4. We perform qualitative and quantitative ***evaluation*** of the proposed method with datasets from two different real-world domains.

The remainder of this paper is organized as follows. In Section 2, we formally define the CTC detection problem. Section 3 presents the related work. Section 4 describes the technical approach, followed by an experimental evaluation in Section 5. Section 6 closes with concluding remarks and directions for future work. See [7] for an extended version of this paper that contains additional details not covered here due to space limitation.

2 Problem Formulation & Definitions

In this section we formally define the problem of contextual time series change (CTC) detection, which can be used on datasets with the following properties:

1. The input data is real-valued, i.e., we do not consider binary or symbolic data.

2. All objects are observed at the same time steps. The time steps do not need to be regularly spaced.

3. Although the data volume may be large, we assume that new data is arriving at a rate such that it can be stored off-line for subsequent processing, i.e., the one-pass streaming data setting is not considered in the current problem formulation. In particular, it is explicitly assumed that all objects and observations seen until the current time step are readily accessible.

We define the following notation: $O = \{o_1, o_2, \cdots\}$ is a target time series; $X = \{x_1, x_2, \cdots\}$ is any time series in the dataset except O; $T = \{t_{i_1}, t_{i_2}, \cdots\}$ indicates a time interval; and O_T and X_T are subsequences of O and X during time period T, respectively.

Generally speaking, an object changes contextually if its *behavior* changes in a different way compared with its *context.* The behavior here is characterized by a (univariate or multivariate) time series, which records the evolution of some property of the object over time. The context is defined by other "similar" time series in CTC detection. We consider both time series O and group G as stochastic, and T_c and T_s as parameters. We call G the *dynamic peer group* of O, T_c the *context construction period*, and T_s the *scoring period*. In definition 1 below we make a statement about the joint measure induced by O and G under different parametric conditions. This is essentially a way of defining the dynamic grouping, where in the context construction period O stays in G almost surely, but may not stay within G in the scoring period. Formally,

DEFINITION 1. *(Contextual Change) A time series O changes contextually at time t if and only if the following two conditions are met.*

1. *There exists a group of time series (G) and a time interval $T_c = \{t-k_1, t-k_1+1, \ldots, t-1\}$ for which*

$$p(O_{T_c} \in G_{T_c}) = 1$$

2. *In the time interval $T_s = \{t+1, t+2, \cdots, t+k_2\}$*

$$p(O_{T_s} \in G_{T_s}) < \epsilon$$

where,$p(O_{T_c} \in G_{T_c})$ and $p(O_{T_s} \in G_{T_s})$ are the probability of O belonging to G in T_c and T_s, respectively. ϵ is a user-defined threshold.

To simplify the problem, we define G in such a way that $p(O \in G)_{T_c} \approx 1$. Since for kNN-based methods it is hard to control the quality of the group members, a threshold-based definition is used.

DEFINITION 2. *(Dynamic Peer Group) The group of time series $\{X^1, \cdots, X^m\}$ constitutes a dynamic peer group G for a time series O in a time interval T_c if and only if for all $j \in (1, m)$*

$$dist(X^j_{T_c}, O_{T_c}) < \epsilon_g$$

$dist(X^j_{T_c}, O_{T_c})$ is an arbitrary distance metric that measures the difference between $X^j_{T_c}$ and O_{T_c}. ϵ_g is a user-defined threshold.

3 Related Work

The concept of analyzing the behavior of a target time series in the context of a group of related series has been explored in prior researches [5, 10, 16], particularly for the problems of fraud detection and temporal outlier detection.

Peer Group Analysis (PGA) is an unsupervised fraud detection method proposed by Bolton and Hand [5]. The basic idea behind PGA is to check whether or not a time step in a time series departs from an expected pattern. The expected pattern is defined by values of the peer group time series at that time step. Peer group of a time series is defined as the k nearest neighbors of the time series using the first n time steps ($n \geq 1$). Since the goal of PGA is primarily fraud detection, these methods focus on scoring a single time step in the detection period[1] [5, 16]. There are two major differences between our approach and PGA. First, in PGA the peer group of a time series is unchangeable throughout the analysis. Whereas in our scheme, peer group is constructed dynamically at each time step, which allows our scheme to handle cases where the peer group of a time series changes with time. Second, because PGA methods focus on scoring a single time step in the detection period, they are not as effective for detecting changes that are more reliably observed over multiple consecutive time steps, especially in noisy data.

Temporal Outlier Detection (TOD) [10] also focuses on detecting outliers within the context of other time series in the dataset. Instead of constructing a peer group for each time series (as done in our approaches), TOD maintains a temporal neighborhood vector that records historical similarities of the target time series to all other time series in the dataset. The outlier score of the target time series at time step t is then given by the L_1 distance between the temporal neighborhood vector at time t and time $t-1$. Given the difference in the way that the target time series is compared with other time series, TOD and our method are meant for finding entirely different types of patterns. For example, if two sets of time series have similar behavior over time within each group, but their behavior as a group starts to differ at a specific time, TOD may flag all these time series as outliers, whereas, our scheme will consider all of these time series as unchanged [7].

4 The proposed approach

In this section, we present the details of our proposed approach for contextual time series change (CTC) detection. The time series to be examined is called the target time series. The proposed method is applied exhaustively in every time step of every time series in the dataset. A change score matrix is reported as the final output. The elements in the matrix indicate how much a specific time series has changed contextually in a given

[1] In some schemes [5], original time series is transformed such that the value at each new time step is a function of several preceding time steps. Then, each time step of this transformed time series is scored individually, even though the score can be a function of multiple time steps of the original time series.

time step. A time series can be labeled as changed if its score is larger than a threshold or its rank is smaller than a threshold.

The proposed approach consists of two steps: context construction and scoring. Context construction is used to discover the dynamic peer group (G) for the target time series (O) during the context construction period (T_c), ensuring that $p(O_{T_c} \in G_{T_c}) \approx 1$. In particular, we build G for O in T_c based on a range query method which naturally follows the definition of dynamic peer group (see Definition 2). We propose a function called k^{th} order statistic distance (Section 4.1), which will be used in the range query method. The time series for which G is constructed successfully (satisfying Condition 1 in Definition 1) are candidates for the second step.

The scoring mechanism provides a probabilistic estimate of the extent to which the target time series follows the behavior of its contextual neighbors in the scoring period (T_s). A non-parametric scoring function called time series area depth is used to measure the deviation of O from G in T_s (Section 4.2). This scoring method does not assume any particular distribution of the datasets and can be used in many different domains.

The proposed method assumes that only a small number of time series in the dynamic peer group changes similarly as the target time series (O). For situations where this assumption cannot be met, we provide a heuristic algorithm (Multimode Remover) to remove the contextual neighbors that have similar behavior as O (Section 4.3).

4.1 Context Construction The purpose of context construction is to discover a dynamic peer group G for a target time series O which can satisfy Definition 2. Therefore, instead of attempting to obtain all the contextual neighbors of O, we aim to find enough time series to adequately describe the context and ensure that $p(O_{T_c} \in G_{T_c}) \approx 1$. Although KNN-based methods easily find members in G, it is difficult to control the quality of the dynamic peer group and thus to meet the condition that $p(O_{T_c} \in G_{T_c}) \approx 1$. In this paper, we choose a range query method, which is an intuitive way to find G. The distance metric is the key component. We use Minkowski distances in this paper to ensure all members in G are similar to O in Euclidean space.

The existence of outliers is a major challenge when using Minkowski distances. For example, L_∞ (used by Li et al. [10]) measures the largest distance of all the time steps during T and is thus highly sensitive to outliers. Therefore, many members will incorrectly be removed from G. Although the impact of outliers is not as large as L_∞, outliers affect the L_1 and L_2 distance as well.

Most standard smoothing methods, such as moving average and Savitzky-Golay filter, cannot address this problem. Instead of ignoring the incorrect information, smoothing methods average the value of outliers into several time steps. When the distance is calculated, a fraction of the incorrect information is still used. In order to overcome this problem, we propose a distance function as described below.

Definition 3. *(k^{th} order statistic distance) The k^{th} order statistic distance between two subsequences O_T and X_T ($O_k(O_T, X_T)$) can be calculated by the following steps.*

1. *Compute point-wise distance between O_T and X_T for each of time steps.*
2. *Rank all the time steps of O_T and X_T in the decreasing order according to the point-wise distance.*
3. *Remove the first $k-1$ time steps in O_T and X_T.*
4. *$O_k(O_T, X_T)$ is given by the predefined Minkowski distance using the remaining time steps.*

The only parameter k affects the performance of the proposed context construction method in two different ways: Many time series are incorrectly included in G when k is too large, while many real members in G are excluded when k is too small. In other words, when k is smaller, the false positive rate of G (FPR_G) is lower but the true positive rate of G (TPR_G) is larger. Thus, the performance of the proposed method is dependent on the choice of k.

In this paper, we provide a method to estimate the best k assuming that the probability of a time step is an outlier can be estimated (in the extended version [7], we propose a supervised method based on cluster sampling theory to estimate this probability). Specifically, we propose a method to choose k which can minimize the FPR_G given the condition that TPR_G should be larger than a user-defined threshold.

Assume that the observed time series X and O can be modeled as

$$x_i = \hat{x_i} + b_x \cdot ol_x + er$$

$$o_i = \hat{o_i} + b_o \cdot ol_o + er$$

where $\hat{x_i}$ and $\hat{o_i}$ are the true values of x_i and o_i. er is weak white noise. ol_x and ol_o are the values of outliers. b_x and b_o are random variables indicating whether or not a time step is an outlier (1 indicates an outlier, 0 otherwise); they follow Bernoulli(p_x) and Bernoulli(p_o), respectively.

LEMMA 1. *Estimation of TPR is given by*

$$\sum_{n=1}^{k-1} \binom{N}{n} q^{N-n}(1-q)^n$$

where N is the number of time steps in T and $q = (1-p_x)(1-p_o)$. Proof is available in [7].

Therefore, k is given by the smallest k which satisfies

$$\sum_{n=1}^{k-1} \binom{N}{n} q^{N-n}(1-q)^n > th_p$$

4.2 Scoring Mechanism Change scoring is the second major step in our CTC detection framework. In this section, we introduce a new robust scoring mechanism, Time Series Area Depth (TAD), to measure the deviation of a target object (O) from its dynamic peer group (G). Fundamentally, TAD is a scoring mechanism derived from statistical depth [17]. Next we briefly introduce the concept of statistical depth and then describe TAD in detail.

Statistical depth measures the position of a given point relative to a data cloud, and evaluates how close the given point is to the center of the data cloud. We prefer a statistical depth function in the proposed method for three reasons. First, they are nonparametric methods. Thus, we do not need to change the method for applications in different domains. Second, they can be used in multivariate analysis. Although in the work presented in this paper we are focused on univariate analysis, the ability to extend the approach to multivariate analysis is under consideration. Third, the property of *affine invariance* ensures robustness of our methodology with respect to the scale, rotation and location parameters of the underlying probability distribution. In particular, it helps address the issue of unequal variance of the observations at different time steps.

DEFINITION 4. *(Time Series Area Depth) The Time Series Area Depth of a time series O given its dynamic peer group G in $T_s = \{t_{i_1}, \cdots, t_{i_N}\}$ is defined as*

$$TAD(T_s) = \sum_{i=i_1}^{i_N} \frac{min(|o_i - c_i^t|, |o_i - c_i^b|)}{\sqrt{|c_i^t - c_i^b|}}$$

where c_i^t (c_i^b) is the value of the top (bottom) m^{th} percentile contour of G at t_i.

m is the parameter to be chosen when using TAD. Here, we choose $m = 68$ in general because it corresponds to one standard deviation if the dataset follows normal distribution.

LEMMA 2. *When G is given, $TAD(T_s)$ is inversely related to $p(O_{T_s} \in G_{T_s})$ under the following assumptions:*

1. *If a time series O belongs to G, for any $i \in \{i_1, i_2, ...\}$*
$$o_i = f_i(G) + \delta_i$$
where $f_i(G)$ is an arbitrary function of all the time series in G and $\Delta = \{\delta_1, \delta_2, \cdots\}$ is a pure random process.

2. *The probability for a given time series O still belongs to the dynamic peer group G at time $t_i \in T_s$ is*
$$\exp\left(1 - \frac{min(|o_i - c_i^t|, |o_i - c_i^b|)}{\sqrt{|c_i^t - c_i^b|}}\right)$$
where all the notations are same as Definition 4. Proof is available in [7].

Next, we will show that the two assumptions are generally true. First, the temporal trend of O is similar to the major trend of G that is constructed by the proposed method. Since $f_i(G)$ is defined as a function that represents the major trend of G at time step t_i, $f_i(G)$ can also be used as the trend of O. Considering that the temporal trend is often responsible for most of the temporal correlation in the time series, all the time steps in Δ are nearly independent to each other (which is the first assumption). The probability given in the second assumption is actually a variant of the Simplical Volume Depth (SVD). The properties of a statistical data-depth function are easily checked for this variant.

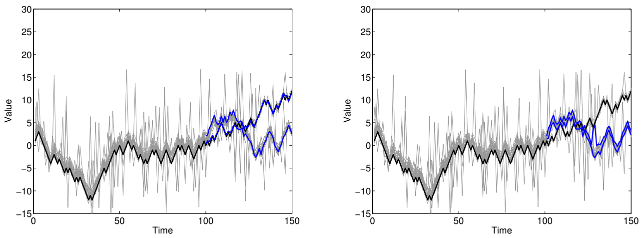
In summary, the advantages to use TAD include:

1. TAD is inversely related to $p(O_{T_s} \in G_{T_s})$ under certain assumptions (See Lemma 2).

2. TAD, although used in univariate analysis in this paper, can be easily expanded to multivariate datasets.

3. Experiments suggest the TAD is parsimonious, in the sense that it performs quite well even if the number of contextual neighbors available is not larger than the dimension of the data (i.e., the number of time steps in the scoring period).

4.3 Dealing with Multiple Modes In some cases, contextual neighbors may change similarly to the target time series. Thus, the dynamic peer group may exhibit multimodal behavior after the time of change. In this situation, scores given by TAD will tend to underestimate the separation. Figure 3a shows an example of this problem wherein 5 out 20 time series (gray) in the



(a) Without Multimode Remover algorithm.

(b) With Multimode Remover algorithm.

Figure 3: A target time series (black) together with its dynamic peer group G (gray). The blue lines are the 68^{th} percentile contour of G after $t = 100$ found without (a) and with (b) Multimode Remover algorithm.

dynamic peer group change similarly as the target time series (black) at $t = 100$. Thus, the 68^{th} percentile contour (the upper and lower bounds of the contour are shown as blues lines) is misleading.

One potential method to address this problem is to find the major mode (we assume that the behaviors of the majority time series in G are similar to each other afterwards). We propose an iterative algorithm called Multimode Remover to obtain the m^{th} percentile contour for the largest mode. Essentially, we first remove the extreme 10% of time series based on their distance to the center of G for each time step. This procedure is iteratively performed until the mean value of G stabilizes. The detailed method is listed in Algorithm 1. Figure 3b shows the 68^{th} percentile contour of the major mode (blue lines) reported by Multimode Remover for the same data used in Figure 3a. Comparing the two figures, we see that the Multimode Remover algorithm successfully identifies the major mode.

Algorithm 1 Multimode Remover

Input: $G = \{X^{j_1} \cdots X^{j_2}\}$ and T_s.
Output: C (m^{th} percentile contour of the major mode in G in T_s).

```
for every t_i ∈ T_s do
    G_i contains the i^th time step of all the members in G.
    while 1 do
        Calculate μ (the mean of G_i).
        Remove 10% values in G_i whose distance are largest.
        Calculate μ' (the mean of the "new" G_i).
        if |μ' − μ| < th then
            STOP;
        end if
    end while
    c_i = m percentile contour in G_i.
end for
```

---

5 Experimental Results

In this section, we demonstrate the capabilities of the proposed CTC detection algorithm on a variety of real-world datasets from the financial and Earth science domains.

5.1 Event Detection based on Historical Stock Market Data

We start with a case study using historical stock market data. In particular, we consider the weekly closing stock prices of S&P 500 companies over the 10-year period from January 2000 through December 2009. This data is publicly available from the Yahoo! Finance website. Since individual stock prices differ in scale and exhibit significant variance, we normalize the raw time series (R) such that

$$x_i = \frac{r_i - min(R)}{max(R) - min(R)}$$

where r_i is the stock price at time t_i, and $max(R)$ and $min(R)$ are the maximum and minimum values in R, respectively.

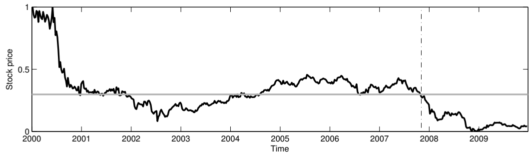
In this experiment, we first show the differences between the contextual changes detected by the proposed method and the traditional changes detected by CUSUM using a case study of Pinnacle West Capital Corporation. Then, we briefly discuss the performance of the proposed CTC detection method on the historical stock market data. The parameters used in this experiment is as below. $T_c = 100$, $T_s = 100$ and $k = 1$.

5.1.1 Case Study of Pinnacle West Capital Corporation

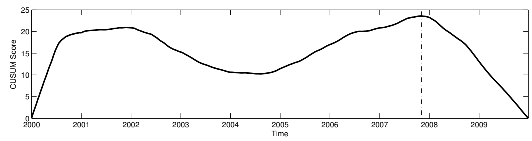
To show the difference between contextual changes and traditional changes detected in the stock market data, we compare the performance of our proposed method with CUSUM[2] using the weekly closing stock prices of Pinnacle West Capital Corporation (Symbol: PNW).

Figure 4 shows the time series of the stock price together with its mean, CUSUM score, and its dynamic peer group constructed by our proposed method. CUSUM detects the point around which the mean of the time series shifts. For this specific example, the changes detected by CUSUM are around the starting point of the global financial crisis from 2007 to 2012. Instead of considering whether or not the pattern of the current subsequence differs from its history, CTC detects changes based on the behavior of other "similar" time series. We note that, while the contextual series continue to rise throughout the latter period (the scoring

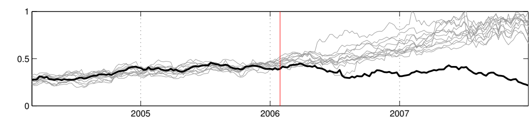
[2] Here, we use CUSUM algorithm prorived by Barnard [2] and choose the target value as the mean of the time series.



(a) The normalized time series of stock price (black) together with its mean (the horizontal gray line).



(b) CUSUM scores (black) and the change point detected by CUSUM (the vertical black dash line).



(c) The normalized time series of PNW price (black) along with its dynamic peer group (gray) that is constructed during the construction period (T_c) using the proposed method.

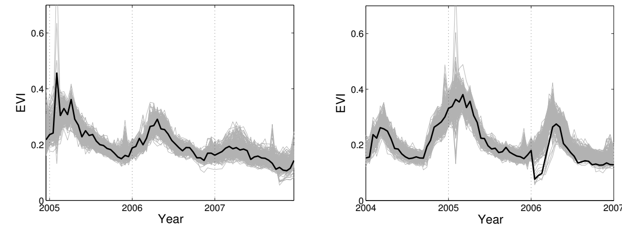
Figure 4: The performance of CUSUM and CTC detection in historical weekly closing stock price of PNW.

period), the PNW stock levels off and suffers some decline. This contextual change happens in 2006, the year in which the company was hit by an \$8 million regulatory setback and the absence of \$7 million in income tax credits it recorded in 2006.

Although PNW belongs to the Electric Utilities sector, most members in its dynamic peer group belong to Stores sector and Processed and Package Goods sector. In fact, none of the members in the dynamic peer group belongs to Electric Utilities sector. In other words, the behavior of PNW is not similar to other stocks in its own sector. One consequence of this is that if its peer group was constructed only based on domain knowledge, this change might not have been detected.

5.1.2 Performance of the Proposed Method in CTC Detection To evaluate the performance of our proposed method in this experiment, we plot the top 50 events reported and visually examine whether or not there is a reasonable separation between the target time series and its dynamic peer group after the time of change. In [7], we provide the time series of the top 10 stocks together with their context.

Generally, all the top 50 stocks reported by the proposed method have reasonable separation compared with their own dynamic peer group. However, the time of change detected by the proposed method for some stock data is not accurate. The main reason is that



(a) Drought in 2007 **(b)** Fire in 2006

Figure 5: Prototypical EVI signals for drought and forest fire events.

many stock prices change gradually – it is difficult to see the separation in the beginning of the change. On the other hand, it is hard to build the dynamic peer group once the gradual change has begun. Thus, there is no clear separation between the target time series and its dynamic peer group shortly after the time of change.

5.2 Forest Fire Detection from Remote Sensing Data In this experiment, we compare the performance of the proposed method and V2DELTA, a traditional change detection method, in forest fire detection. We draw two conclusions from this experiment. First, the contextual change is a distinguishing feature that can be used to discover fires against droughts. Second, our proposed method is capable of successfully detecting contextual changes in EVI datasets. Detailed descriptions of the dataset and validation data and the discussion related to false positives of the proposed method can be found in [7]. The parameters used in this experiment are: $T_c = 46$, $T_s = 6$ and $k = 5$.

5.2.1 Limitation of traditional time series change detection The Enhanced Vegetation Index (EVI), an indicator of "greenness" reflected from the earth's surface, is used as the input dataset. EVI is one of the most widely used signals for forest fire detection [12]. However, many other events, such as drought, can also cause a decrease in EVI signals. Although many attempts have been made to increase the accuracy of forest fire detection using traditional change detection methods, detecting fires in the context of drought based on EVI signals is still an open problem faced with the following challenges:

- There are many different land cover types in this region. Hence, the drops due to drought are not necessarily smaller than the drops due to fire.
- The quality of data is not always good due to obstruction from smoke or atmospheric interference.
- No definitive set of distinguishing features based on

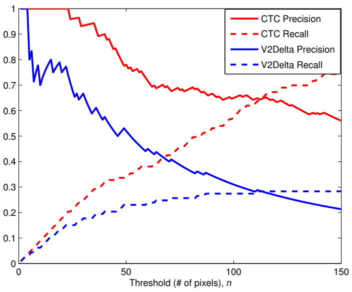


Figure 6: Precision and recall curve of the proposed algorithm and V2DELTA in the experimental region.

EVI has been discovered to be used in fire detection against drought.

Figure 5 shows examples of EVI signals under drought and fire, respectively. Because drought is an event under which many time series exhibit a change, as shown in Figure 5a, it typically does not lead to a CTC in the time series. However, fires generally affect limited regions and thus can be detected as a CTC in EVI, as shown in Figure 5b.

5.2.2 Comparison with V2DELTA

We compare the proposed algorithm with V2DELTA, which is designed to detect land cover changes based on EVI [13]. V2DELTA is a traditional change detection algorithm, meaning that it learns a model from a portion of the input time series while normalizing for the historical variability, then makes a prediction for some window and measures whether or not the observed data deviates from that prediction. As a result, V2DELTA reports all kinds of events, including both fires and droughts, which lead to deviations of the observed time steps from its own historical data.

Figure 6 (best viewed in color) shows the precision-recall curves [15] for the proposed algorithm (red lines) and V2DELTA (blue lines) in the experimental region. From this result, we note that a significant improvement in both precision and recall is achieved for the proposed algorithm. Figure 7 and Figure 8 show two examples to illustrate this observation.

6 Conclusions & Future Work

In this paper, we presented a framework for a class of time series analysis problems called *contextual time series change* (CTC) detection, and we proposed a two-step algorithm to address it. The novelty of the proposed algorithm includes robustness to common data characteristics such as outliers (the k_{th} order statistic distance), multimodal behavior (Multimode Remover algorithm) and a new scoring mechanism to measure the deviation of a target time series from the others (time series area depth).

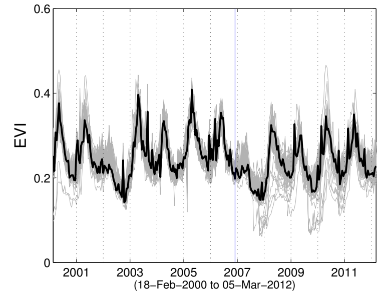


Figure 7: Examples of false positive of V2DELTA caused by the drought in 2007 which is detected as normal objects by the proposed algorithm. Compared with the data from 2003 to 2006, the drop in EVI in 2007 is obvious. Therefore, V2DELTA detect an event happened in 2007 in the target time series (the black line). However, the drop in 2007 is a common feature of all the time series in the contextual group (gray). Thus, the proposed algorithm gives a relatively small score to the target time series.

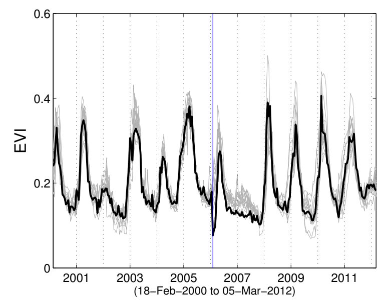


Figure 8: Examples of false negative of V2DELTA which is detected as positive objects by the proposed algorithm. The target time series (black) shows different behavior from the context (gray) after the time of change (blue), which indicates that it is a CTC. The proposed method detects it correctly as a CTC change. However, the magnitude of change in this time series is not large enough to be labeled as traditional changes by V2DELTA.

We provide a theoretical proof of the optimized performance of the k_{th} order statistic distance under certain assumptions and the derivation that TAD is inversely correlated to changes in the probability of a target object belonging to a certain context. This inverse relationship of TAD ensures that it is a good indicator of CTC. We also show other good properties of TAD, for example, it does not depend on the dataset, it does not require a large amount of contextual neighbors, it is superior to naïve approaches, and it can be easily expanded to multivariate analysis.

Two real datasets from the financial and Earth science domains have been used to demonstrate the unique capabilities of the proposed algorithm. From the experimental results, we notice that CTC generally indicates a new type of events compared with the results of traditional time series change detection, such as CUSUM and V2DELTA. In particular, two experiments were performed. We first compared the proposed method with CUSUM in the weekly closing stock prices of S&P 500 companies. The events detected as contextual changes generally related to internal events at the affected companies, e.g., release of lower forecasts or changes in management. CUSUM, instead of detecting such contextual changes, reports changes in the overall financial market, e.g., global recessions. The second experiment used the vegetation time series dataset (EVI). We compared the results of CTC detection with V2DELTA, which is designed to detect land cover changes. From this experiment, we conclude that unlike EVI, which detects all types of changes, CTC is capable of identifying sub-area events. Both the quantitative result using precision and recall as well as real examples have shown to support this conclusion.

There are several open challenges that remain in CTC detection. The hidden relationship, which defines the contextual time series group, is defined in a somewhat naïve way in this paper. As a result, many CTC's are not detected because of an inadequate number of contextual neighbors. The framework could be extended to accommodate other measures such as correlation and dynamic time warping distance [4]. The proposed methods are also built based on non-parametric methods, which is good when the models of data are unknown. However, in situations where well-defined time series models exist for the data under consideration, one could develop an approach that incorporates the model into the context, potentially significantly improving performance. Besides, the proposed method contains several user defined parameters. Further study related to choosing the best parameters may help achieve better performance. Finally, the proposed method is a brute-force method. Thus, there are several improvements that can increase efficiency; specifically, when the input time series has temporal autocorrelation, this property can potentially be exploited to reduce the number of similarity computations.

Acknowledgment

This research was supported in part by the National Science Foundation under Grants IIS-1029711 and IIS-0905581, as well as the Planetary Skin Institute. Access to computing facilities was provided by the University of Minnesota Supercomputing Institute.

References

[1] T. W. Anderson. *The Statistical Analysis of Time Series*. John Wiley & Sons, 1994.

[2] G. Barnard. Control charts and stochastic processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 239–271, 1959.

[3] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, 1993.

[4] D. J. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In *Proceedings of KDD '94: AAAI Workshop on Knowledge Discovery in Databases*, pages 359–370, 1994.

[5] R. Bolton and D. Hand. Unsupervised profiling methods for fraud detection. In *Credit Scoring and Credit Control VII*, 2001.

[6] J. Chen and A. Gupta. On change point detection and estimation. *Communications in Statistics: Simulation & Computation*, 30(3):665–697, 2001.

[7] X. C. Chen, K. Steinhaeuser, S. Boriah, S. Chatterjee, and V. Kumar. Contextual time series change detection. Technical Report 12-018, Computer Science, University of Minnesota, 2012.

[8] V. Guralnik and J. Srivastava. Event detection from time series data. In *KDD '99: Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 33–42, 1999.

[9] T. L. Lai. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(4): 613–658, 1995.

[10] X. Li, Z. Li, J. Han, and J.-G. Lee. Temporal outlier detection in vehicle traffic data. In *ICDE '09: Proceedings of the IEEE International Conference on Data Engineering*, pages 1319–1322, 2009.

[11] T. W. Liao. Clustering of time series data—a survey. *Pattern Recognition*, 38(11):1857–1874, 2005.

[12] V. Mithal, A. Garg, S. Boriah, M. Steinbach, V. Kumar, C. Potter, S. Klooster, and J. C. Castilla-Rubio. Monitoring global forest cover using data mining. *ACM Transactions on Intelligent Systems and Technology*, 2: 36:1–36:24, 2011.

[13] V. Mithal, A. Garg, I. Brugere, S. Boriah, V. Kumar, M. Steinbach, C. Potter, and S. A. Klooster. Incorporating natural variation into time series-based land cover change detection. In *Proc. Conference on Intelligent Data Understanding (CIDU)*, pages 45–59, 2011.

[14] A. Silvestrini and D. Veredas. Temporal aggregation of univariate and multivariate time series models: A survey. *Journal of Economic Surveys*, 22(3):458–497, 2008.

[15] P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison-Wesley, Boston, MA, 2006.

[16] D. Weston, D. Hand, N. Adams, C. Whitrow, and P. Juszczak. Plastic card fraud detection using peer group analysis. *Advances in Data Analysis and Classification*, 2:45–62, 2008.

[17] Y. Zuo and R. Serfling. General notions of statistical depth function. *The Annals of Statistics*, 28(2):461–482, 2000.